

A flexible multivariate random effects proportional odds model with application to adverse effects during radiation therapy

Nicole Augustin¹, Sung Won Kim^{2,3*},

Annemarie Uhlig⁴, Christina Hanser⁴, Michael Henke⁴, Martin Schumacher²

¹ Department of Mathematical Sciences, University of Bath, Bath, UK

² Institute of Medical Biometry and Statistics, University of Freiburg

³ Freiburg Center for Data Analysis and Modeling, University of Freiburg

⁴ Section Clinical Studies, Department of Radiation Oncology,
University Hospital Freiburg

February 29, 2016

Abstract

Radiation therapy in patients with head and neck cancer has a toxic effect on mucosa, the soft tissue in and around the mouth. Hence mucositis is a serious common side effect and is a condition characterized by pain and inflammation of the surface of the mucosa. Although the mucosa recovers during breaks of and following the radiotherapy course the recovery will depend on the type of tissue involved and on its location. Statistical models used in oncology for analysing oral mucositis are often too simplistic and there is a need for improved statistical methods in order to help improve planning of radiation therapy. We

*contributed equally to this work.

present a novel flexible multivariate random effects proportional odds model which takes account of the longitudinal course of oral mucositis at different mouth sites and of the radiation dosage (in terms of cumulative dose). The model is an extension of the *proportional odds model* which is used for ordinal response variables. Our model includes the ordinal multivariate response of the mucositis score by location, random intercepts for individuals and includes a non-linear function of cumulative radiation dose. The model allows to test whether sensitivity differs by mouth sites after having adjusted for site specific cumulative radiation dose. The model also allows to check whether and how the (non-linear) effect of site specific dose differs by site. We fit the model to longitudinal patient data from a prospective observation and find that after adjusting for cumulative dose, upper, lower lips and mouth floor are associated with the lowest mucositis scores and hard and soft palate are associated with the highest mucositis scores. This implies the possibility that tissues at different mouth locations differ in their sensitivity to the toxic effect of radiation. We also find that cumulative dose followed by mouth site are the strongest predictors of mucositis, and the effects of age and gender are negligible.

Keywords: generalised additive mixed model, proportional odds model, cumulative threshold model, radiation therapy, multivariate repeated measures, ordered categorical response, biologically effective dose (BED), mucositis.

1 Introduction

Radiation therapy in patients with head and neck cancer has a toxic effect on mucosa, the soft tissue in and around the mouth. Hence mucositis is a common side effect and is a condition characterized by pain and inflammation of the surface of the mucosa. Mucositis causes serious oral pain which makes the food intake of the cancer patients painful. It is known that mucositis is in general developed 14 days after starting radiotherapy. Although the mucosa recovers during breaks of and following the radiotherapy course the recovery will depend on the type of tissue involved and on its location. Hence there is a need to quantify the damage to the mucosa by different mouth sites. Here the clinical objective is in assessing the sensitivity to radiation of different mouth sites in order to improve planning of radiation therapy to avoid these side effects.

To obtain clinical evidence observational data were collected on a total of 75 patients receiving radiation therapy between September 2006 and February 2012. During the course of radiation therapy mucositis was assessed in 8 mouth sites approximately twice a week. For assessment an ordinal severity score with 5 levels was used, ranging from 0 (no change in tissue) to 4 (necrosis). At each assessment cumulative dose and the median value of the percentage site specific dose was recorded, so that the site specific cumulative dose could be derived. In addition all relevant patient characteristics were recorded.

Statistical analyses on oral mucositis as an adverse effect of therapy are often simplistic. For example Sakellari et al. [13] analyse oral mucositis in patients after high-dose chemotherapy using logistic regression of cumulative incidence of mucositis. The approach neglects the time course of the disease (oral mucositis), the fact the end-point is ordinal and the dosage of the chemotherapy.

Our aim is to answer the clinical question by modelling the outcome variable mucositis score as a function of patient specific variables and (site) specific cumulative dose. Then the sensitivity of different mouth sites to radiation therapy can be assessed by estimating the effect of mouth sites on the mucositis score, after adjusting for confounder variables, the most obvious being site specific cumulative dose. Due to the observational nature of the data we need to adjust for confounder variables.

These longitudinal data on the course of mucositis and cumulative dose of radiation therapy have a complex structure and pose a number of statistical challenges:

(1) the outcome is ordinal; (2) We have a multivariate outcome variable, as we have mucositis scores assessed for eight different sites per patient and time points (evaluation); (3) the outcome is longitudinal and we expect that repeated observations of individuals are correlated; (4) due to the fact that the mucosa can recover during radiation breaks a non-linear effect of sites specific cumulative dose on the mucositis score is plausible. (5) Finally, it is also likely that the effect of site specific dose differs by site, i.e. there is an interaction between site and site specific dose. We address these challenges with a flexible multivariate random effects proportional odds model. This is an extension of the *proportional odds model* [11] which is used for ordinal response variables and belongs to the class of generalized linear models used for modelling

the dependence of an ordinal response on discrete or continuous covariates. The model is an extension of the multivariate random effects models for longitudinal data described in Verbeke et al. [15]. See also Faraway [7] and Agresti [1] for some recent expositions of ordinal response models.

2 The data

Observational data were collected on a total of 75 patients receiving radiation therapy for head and neck cancer between September 2006 and February 2012 in the Radiation Oncology Department, Medical Center, University of Freiburg. Three-dimensional radiation planning techniques and standard fractionation (5 x 2.0 Gy/wk) were followed. The total dose to the tumour was 60 Gy to 70 Gy. Radiation breaks were not allowed. Additional cisplatin 100 mg/m² was given intravenously on days 1 and 22 (and on day 43 for patients undergoing primary definitive radiation or with residual cancer following surgery). The oral mucosa was evaluated in eight mouth sites (upper lip, lower lip, floor of mouth, right cheek, left cheek, tongue, soft palate hard palate) by using the five level NIH-CTC oral toxicity scale ((0) no mucositis, (1) erythema, (2) ulcer, (3) confluent ulcer, (4) necrosis). Twice-weekly assessments (at least 3 days apart) continued throughout radiotherapy and then until resolution of oral mucositis. At each assessment the cumulative dose to the tumor and the median value of the percentage of the site specific dose was recorded, so that the site specific cumulative dose could be derived. In addition all relevant patient characteristics were recorded. All study procedures were reviewed by the local ethics committee and all patients provided written consent.

The mucositis progressions on eight different sites of four randomly sampled patients are depicted in Figure 1, showing the ordinal scale of mucositis scores as well as longitudinal multivariate structure.

Our main goal is to investigate the relationship between mucositis progression and potential risk factors. Let y_{ijl} be the mucositis score for individual $i = 1, \dots, 75$, evaluation $j = 1, \dots, n_i$ and site $l = 1, \dots, 8$. This gives a total number of 7314 observations (75 patients \times 8 sites \times 12.23 the mean number of evaluations per patient). The risk factors are patient characteristics

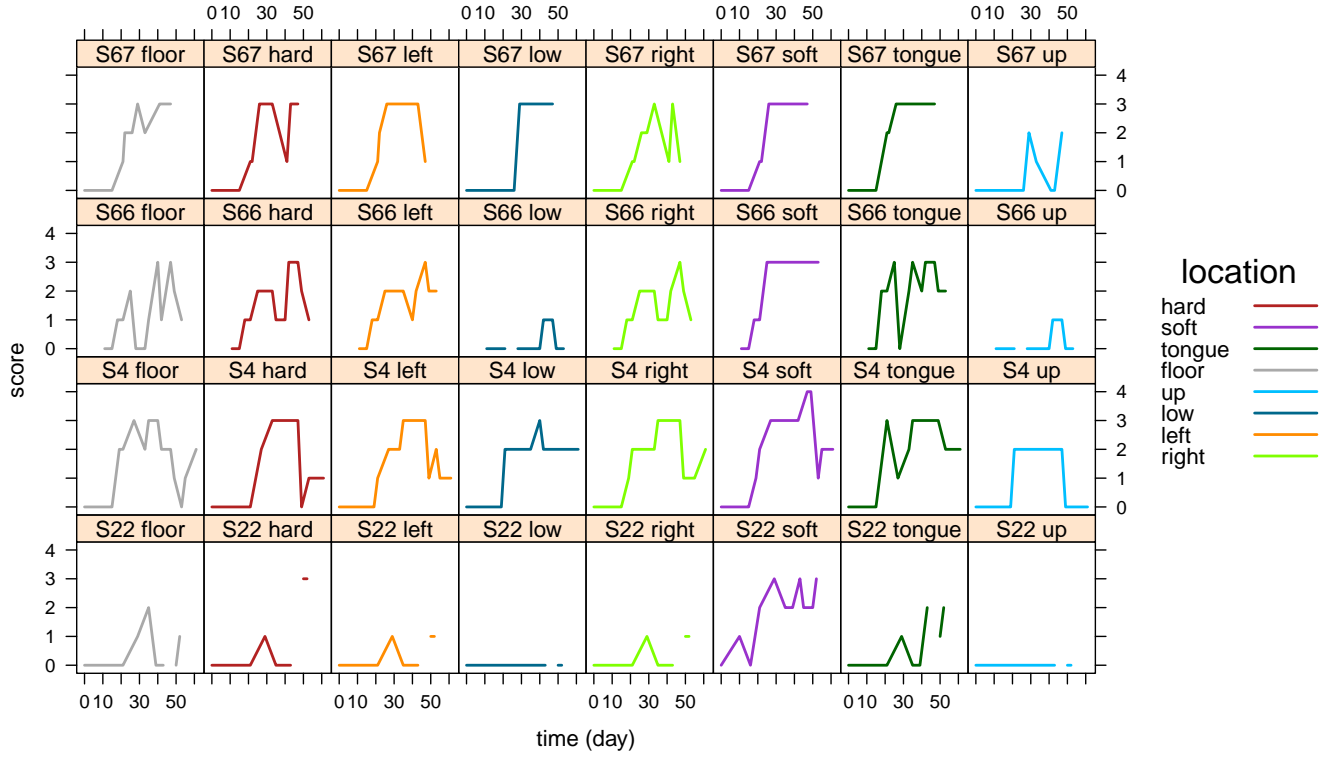


Figure 1: Each row shows mucositis progression for one of four randomly selected individuals on 8 mouth sites (hard palate, soft palate, tongue, mouth floor, upper lip, lower lip, left cheek, right cheek); the same color indicates the same site.

(sex_i and age_i), site_l (hard palate, soft palate, tongue, mouth floor, upper lip, lower lip, left cheek, right cheek), cumulative (radiation) dose (cumdose_{ij}), median percentage dose of the site (perc_{il}), the site specific cumulative dose cumdos.site_{ijl} and the site specific volume_{il} . The site specific dose cumdos.site_{ijl} at evaluation j is determined by multiplying the total cumulative dose cumdose_{ij} with the median percentage dose of the site (perc_{il}). The median percentage dose per site (perc_{il}) is assumed to be constant over time as it is the median percentage of cumdose_{ij} at the site receiving radiation over the entire radiation schedule. Apart for some exceptions, time in days is proportional to cumulative dose (Figure 3), hence in the following modelling we will use cumulative dose and site specific cumulative dose.

Table 1 shows that no mucositis (0) is observed most often, followed by erythema, ulcer, confluent ulcer and necrosis. Figure 2 shows frequencies of mucositis score by site and shows that

Table 1: Collapsing the four level ordinal mucositis scores into binary scores using different cut points r . Scores are no mucositis (0), erythema (1), ulcer (2), confluent ulcers (3) and necrosis (4). The total number of n observations is due to 75 patients \times 8 sites \times 12.23 the mean number of evaluations per patient.

	mucositis score					
cut point	0	1	2	3	4	n
none	3641	1616	1433	639	10	7339
$r = 0$	3641	3698				7339
$r = 1$	5257		2082			7339
$r = 2$	6690			649		7339
$r = 3$	7329				10	7339

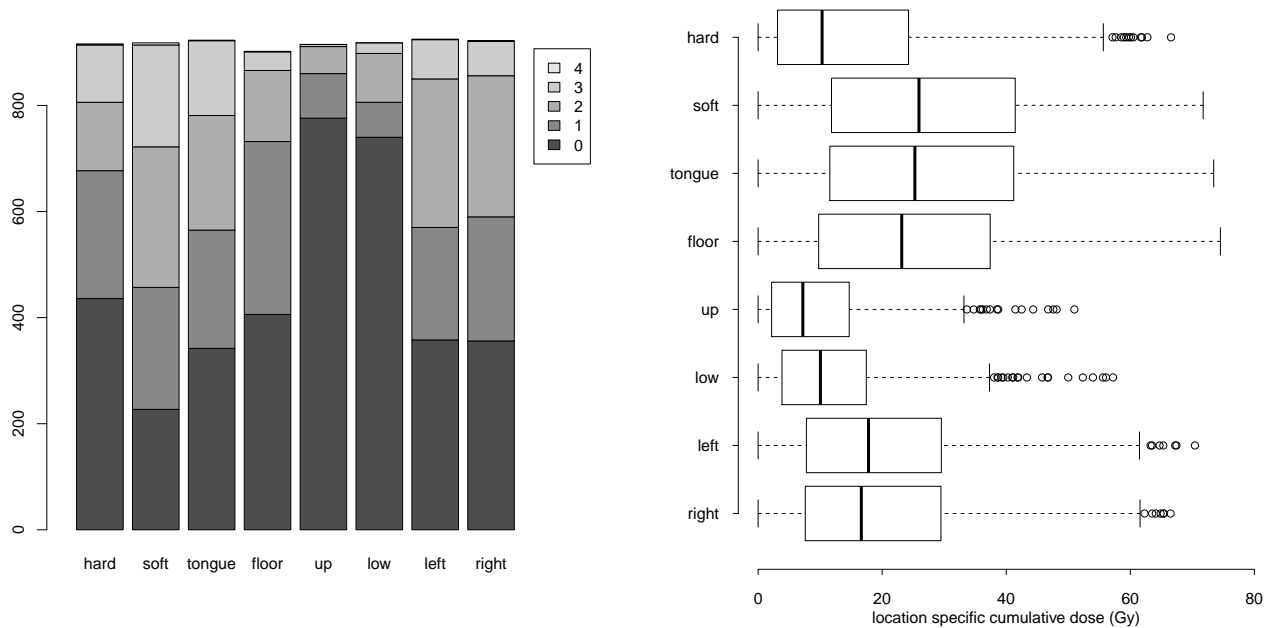


Figure 2: Left: Frequencies of mucositis scores by site. Scores no mucositis (0), erythema (1), ulcer (2), confluent ulcers (3) and necrosis (4). Right: Boxplots of site specific median cumulative dose for mucositis score 0 to 4.

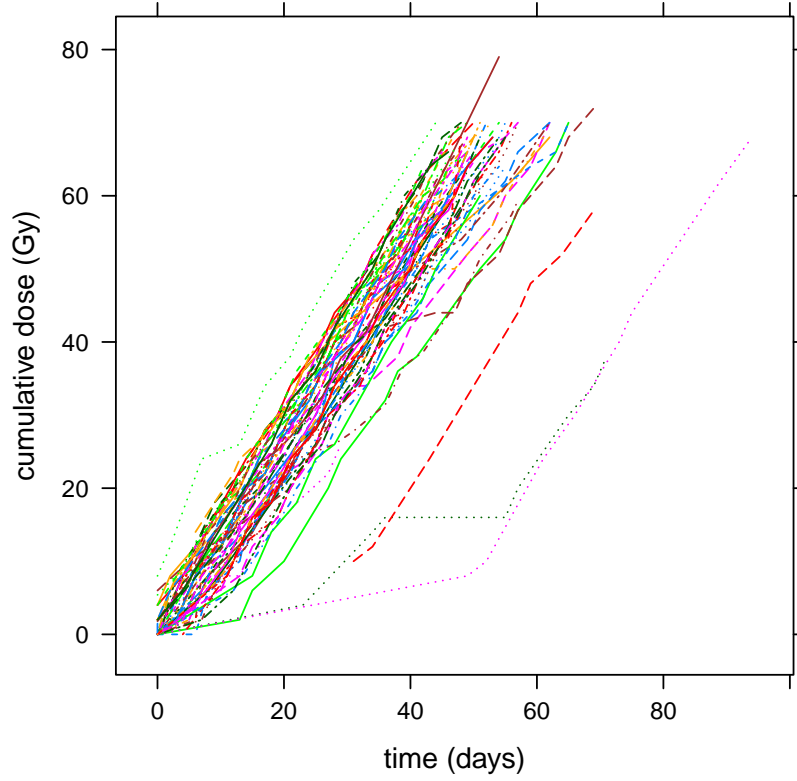


Figure 3: Cumulative dose versus time in days.

on the lower and upper lip high mucositis scores are rarer than at other sites. Figure 2 shows the cumdos.site_{ijl} by site and mucositis score: upper and lower lip tend to receive the lowest levels of site specific dose cumdos.site_{ijl} and mouth floor, tongue and soft palate tend to receive the highest doses.

3 Biologically effective dose (BED)

Here we analyse physically cumulative dose. To eventually correct for repopulation we also calculated the biologically effective dose (BED) using the so called $\alpha\beta$ model, see for a review Fowler [8]. Here we assume $\frac{\alpha}{\beta} = 10\text{Gy}$, where α and β are the coefficients for dose and dose² in a linear model for cell kill. In the case of mucosa the $\alpha = 0.3\text{Gy}^{-1}$, the onset of accelerated repopulation is 7 days and the average doubling time during accelerated repopulation is assumed

to be 2.5.

Figure 4 compares BED versus cumulative dose for each site. The straight lines, except for some kinks, indicate that the two measures are proportional to each other. The kinks originate from breaks in the radiation schedule due to strong side effects for some patients. These breaks affect BED but not the cumulative dose.

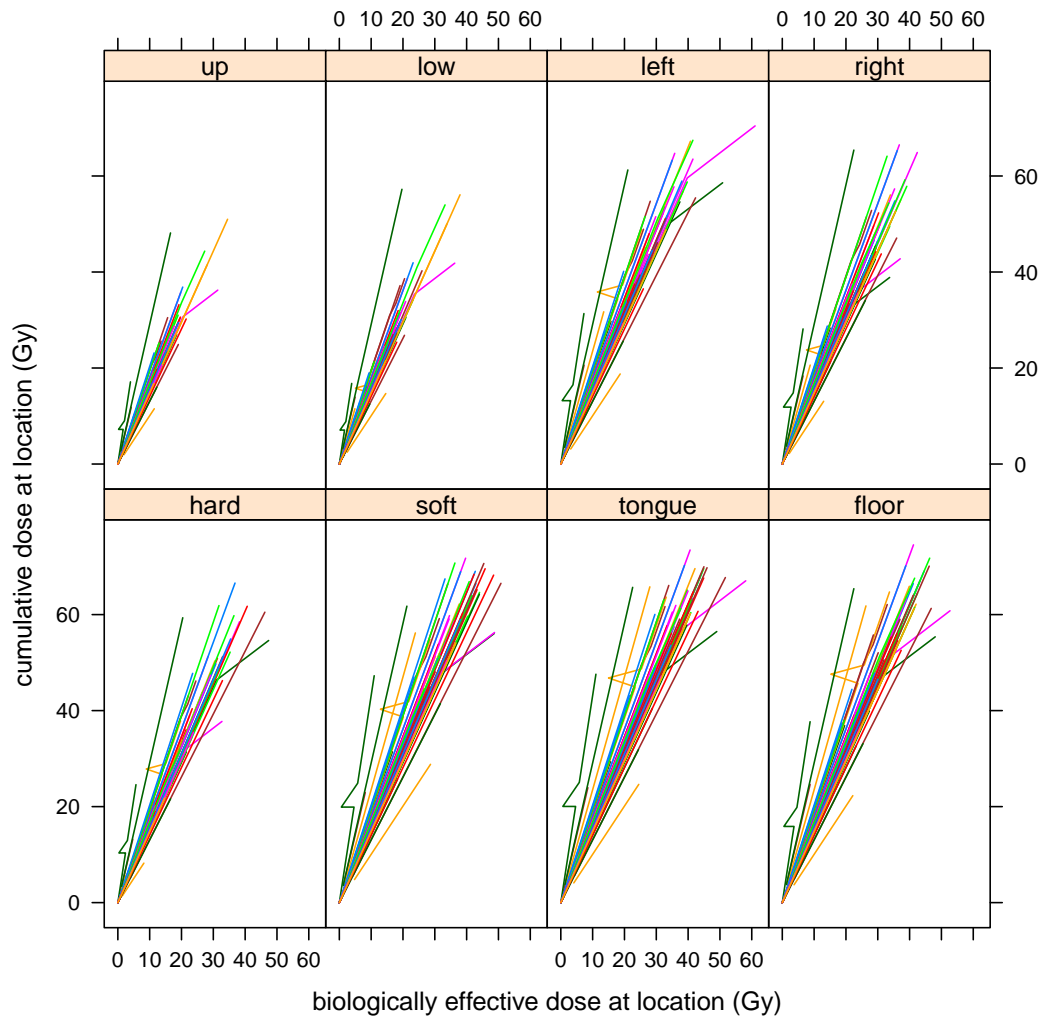


Figure 4: Cumulative dose vs. acute mucosa biological effective dose (BED).

4 The flexible proportional odds model with random effects

As outlined in the introduction due to the complex data situation we require a model for an ordinal multivariate outcome variable. In addition the model should be able to accommodate the correlated repeated measures of individuals and should allow effects of continuous variables to be non-linear, and we also want to test whether there is an interaction between site and site specific dose.

4.1 The proportional odds model

The proportional odds model [11] addresses the first point, as it is a class of generalized linear models used for modelling the dependence of an ordinal response on discrete or continuous covariates. It can also be seen as an extension of a logistic regression model [4], as we could model the mucositis scores collapsed into two categories for the two events $Y \leq r$ and $Y > r$ as shown in Table 1. But this leads to loss of efficiency, resulting in larger standard errors [1].

Let Y_i denote the mucositis severity score in the range 0, ..., 4 for individual i . Then we model the cumulative probability of Y_i being less or equal to mucositis score r , $P(Y_i \leq r)$ on the logit scale. That is on the log scale of the (cumulative) odds of the event $Y_i \leq r$:

$$\text{logit}(P(Y_i \leq r)) = \log \frac{P(Y_i \leq r)}{P(Y_i > r)} = \alpha_r - x_i \beta. \quad (1)$$

On the logit scale the model is linear which makes it computationally convenient. In the above example the x_i is a continuous variable, e.g. cumulative dose. In the case of mucositis a plausible way of constructing model 1 is to use the concept of an unobserved continuous response. Suppose U_i is the unobserved mucositis severity on a continuous scale with $U_i = x_i \beta + \epsilon_i$. Although we only observe the discrete mucositis scores, $Y_i = r$ is only observed if $\alpha_{r-1} < U_i \leq \alpha_r$. The "cut-points" α_r are unknown points on the continuous mucositis severity scale. Lets also assume that $\epsilon_i = U_i - x_i \beta$ has a logistic distribution function $F(\cdot)$ with $F(\cdot) = \frac{\exp(\cdot)}{1 + \exp(\cdot)}$ then:

$$P(Y_i \leq r) = P(U_i \leq \alpha_r) = P(U_i - x_i \beta \leq \alpha_r - x_i \beta) = F(\alpha_r - x_i \beta)$$

and

$$P(Y_i \leq r) = \frac{\exp(\alpha_r - x_i \beta)}{1 + \exp(\alpha_r - x_i \beta)}$$

This model is mathematically equivalent to model 1. We also note the intuitive interpretation, if $\beta > 0$ as x_i increases also the $P(Y_i > r)$ increases. For example, as cumulative dose increases, it is more likely that a high mucositis score is observed. In a model containing a linear effect for `cumdos.site` the odds ratio of the event $Y \leq r$ for $x_1 = \text{cumdos.site} - 10$ and $x_2 = \text{cumdos.site}$, whilst keeping all other variables fixed, is

$$\frac{P(Y_i \leq r|x_1)/P(Y_i > r|x_1)}{P(Y_i \leq r|x_2)/P(Y_i > r|x_2)} = \exp(-(x_1 - x_2)\beta) = \exp((x_2 - x_1)\beta) = \exp(10\beta).$$

When using the same β across all categories, we assume that the odds ratio does not depend on the category r , i.e. these cumulative odds are proportional across all categories. In a model containing a factor variable for site based on the frequencies in Figure 2, this implies that the cumulative odds of the events $Y \leq r$, i.e. $Y \leq 0$, $Y \leq 1$, $Y \leq 2$ and $Y \leq 3$ by site are proportional across all categories. For example, comparing the cumulative odds of mucositis in hard palate versus soft palate is the same for all r :

$$\frac{P(Y_i \leq 0|\text{hard})/P(Y_i > 0|\text{hard})}{P(Y_i \leq 0|\text{soft})/P(Y_i > 0|\text{soft})} = \dots = \frac{P(Y_i \leq 2|\text{hard})/P(Y_i > 2|\text{hard})}{P(Y_i \leq 2|\text{soft})/P(Y_i > 2|\text{soft})}$$

4.2 A flexible multivariate random effects proportional odds model

In order to address all of the model requirements we extend the proportional odds model by incorporating a multivariate response, random effects and non-linear functions of continuous variables. This is an extension of the class of generalised additive mixed models [6, 16], because the logistic distribution we are assuming for the latent response variable U_i and the error ϵ_i does not belong the exponential family. Below we introduce these different features step by step.

The model for individual i , assessment time index j , and site l is specified as

$$\text{logit } P(Y_{ijl} \leq r) = \log \frac{P(Y_{ijl} \leq r)}{P(Y_{ijl} > r)} = \alpha_r - \eta_{ijl}$$

where

$$\text{linear model: } \eta_{ijl} = \mathbf{x}_{ijl}^T \boldsymbol{\beta} + b_i \quad (2)$$

and $U_{ijl} = \eta_{ijl} + \epsilon_{ijl}$ and ϵ_{ijl} follows the logistic distribution. The \mathbf{x}_{ijl} is a row of the model matrix containing the explanatory variables `agei`, `studyi`, `sexi`, `cumdos.siteijl`, `sitei`, where `sitei`

is a vector with dummy variables for the 8 sites and the parameter vector $\boldsymbol{\beta}$ contains the respective coefficients. We deal with the multivariate response, the mucositis scores assessed for eight different mouth sites per patient and evaluation by including the factor site_i . This allows us to test the null hypothesis that the effects of site are equal. As the dose specific to site is determined by the tumour location, we adjust site estimates by site specific dose cumdos.site_{ijl} . Individuals are observed repeatedly in time and within individual observations y_{ijl} are correlated in time and between locations and we account for this by introducing a random intercept b_i for individual i with $b_i \sim N(0, \sigma_b^2)$. As we expect the dose-response relationship not to be linear, we use a flexible function to model the effect of sites specific cumulative dose:

$$\text{reduced model: } \eta_{ijl} = \mathbf{x}_{ijl}^T \boldsymbol{\beta} + f(\text{cumdos.site}_{ijl}) + b_i \quad (3)$$

the smooth function f is represented using a thin plate regression spline [16] with a penalty based on the second derivative of the smooth function and now $\mathbf{x}_{ijl}^T = (\text{age}_i, \text{study}_i, \text{sex}_i, \text{site}_i^T)$. The model above can be further extended. Besides testing whether sites effects differ, we are also interested in testing whether the effect of site specific dose differs by site. For this we introduce an interaction between site and site specific dose, by letting f vary by site_i :

$$\text{full model: } \eta_{ijl} = \mathbf{x}_{ijl}^T \boldsymbol{\beta} + f_i(\text{cumdos.site}_{ijl}) + b_i. \quad (4)$$

4.3 Parameter estimation

Inference can be based on Markov chain Monte Carlo (MCMC) [5, 2], on a mixed model approach using restricted maximum likelihood (REML) [10, 6, 17]. All of these are implemented in the BayesX package [3, 14]. We present estimates using a Laplace approximation to REML of Wood et al. [17], implemented in the `gam()` function of the R ([12]) package `mgcv`. In the supplementary material we supply example R code for model fitting in both BayesX and in `mgcv`.

Although the above reduced model (3) and the full model (4) include additive non-linear predictors and random effects they can be expressed as a penalised generalised linear model (GLM); $\eta_{ijl} = \mathbf{x}_{ijl} \boldsymbol{\theta}$, where \mathbf{x}_{ijl} is a row of the model matrix containing *all* components of the model; that is, all strictly parametric components, such as explanatory variables of linear effects and the

basis functions for non-linear predictors evaluated at observations ijl . The parameter vector θ contains the coefficients of the basis functions and coefficients of linear terms and the cut-points α_1 to α_4 . Note that for identifiability the first cut point α_1 is set to -1.

The parameter estimation is an extended version of the nested iteration scheme described in Wood [18]. Due to the ordered categorical response the scheme also needs to accommodate the variable number of ordered cut points (here α_1 to α_4). The outer iteration is a Laplace approximation marginal (restricted) maximum likelihood (LAML) estimation of smoothness parameters and the cut points, where the coefficients of linear terms, and of basis functions for the smooth terms of the parameter vector θ , are integrated out. The inner iteration is a penalised iterative re-weighted least squares (PIRLS) algorithm to find all other parameters, i.e. the coefficients of basis functions, and coefficients of linear terms. For details see Appendix F in Wood et al. [17].

4.4 Model selection and validation

We consider the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) for model selection. The AIC is defined as $AIC = -2l(\hat{\theta}) + 2edf$, the BIC is $BIC = -2l(\hat{\theta}) + \log(n)edf$, where l is the conditional log likelihood given the penalised parameters, edf are the effective degrees of freedom, here estimated by the trace of the matrix that maps the un-penalised estimates onto the penalised estimates corrected for the smoothing parameter uncertainty as described in Wood et al. [17]; see also Greven and Kneib [9]. We also consider the root mean square prediction error (RMSPE) estimated by cross-validation as a selection criterion. For the cross-validation we leave out in turn each of 15 mutually exclusive sets of patients and fit the model to the remaining patients data and predict the probabilities for the outcomes, the mucositis scores. Having obtained predictions for each set of patients left out we can estimate the root mean prediction error for each mucositis score.

5 Results

In the modelling we combine the mucositis scores 3 (confluent ulcer) and 4 (necrosis) due to the low frequencies of score 4. We start the model selection by reducing the complexity of the full model (4). Table 5 gives an overview of the considered models and the results of selection criteria are shown in Table 5. The most complex model is the full model (4) overall has marginally the best goodness of fit, but with regards to parsimony the reduced model (3) with the terms for age and sex removed comes out best. Hence the final selected model is

$$\text{best model: } \eta_{ijl} = \alpha + \beta_1 \text{study}_i + \text{site}_i^T \beta_2 + f(\text{cumdos.site}_{ijl}) + b_i \quad (5)$$

where β_2 is the parameter vector for the 8 sites. Checking the residuals of the best model (5), shows that the model fits well. The distributional assumptions of the model are correct, i.e. the residuals follow a logistic distribution (Figure 5) and the random effects follow a normal distribution (Figure 6). The assumption of proportional odds is adequate and has been checked using diagnostic plots (not shown).

The model selection results also show that `cumdos.site` is by far the most important predictor, followed by `site`. Results also show that effect of site specific cumulative dose is not linear, as the goodness of fit is substantially reduced when the we use a linear effect for `cumdos.site`. This is also apparent from the estimated smooth function of `cumdos.site` shown in Figure 6. Replacing `cumdos.site` by `BED` in the full model (4), the reduced model (3) and the linear model (2) does not improve the model fit. These models give an AIC of 15028.67, 15079.57 and 15085.05 respectively and the BIC values are 15726.93, 15658.66 and 15639.53. The effect of `cumdos.site` is shown on the scale of the unobserved continuous mucositis score. The cumulative dose has initially, up to approximately 20 Gy quite a strong effect on mucositis, and this effect becomes weaker for > 20 Gy. There is a considerable amount of uncertainty about the effect for high cumulative dosages. Figure 7 shows the cumulative odds of having a score greater than r over having a score less or equal to r , with hard palate as the reference. e.g.:

$$\frac{P(Y_i > r|\text{soft})/P(Y_i \leq r|\text{soft})}{P(Y_i > r|\text{hard})/P(Y_i \leq r|\text{hard})} = \frac{P(Y_i \leq r|\text{hard})/P(Y_i > r|\text{hard})}{P(Y_i \leq r|\text{soft})/P(Y_i > r|\text{soft})} = \exp(0.185) = 1.171.$$

This shows that, after adjustment for cumulative radiation dose, upper, lower lip and mouth

floor are much less likely to have a higher mucositis score than the other sites. The site soft palate is most likely to have a high score.

Table 2: Model overview given in ascending order regarding model complexity. The model equation numbers are given in brackets. The \mathbf{x}_{ijl} is a row of the model matrix containing the explanatory variables age_i , study_i , sex_i , cumdos.site_{ijl} , site_i , where site_i is a vector with dummy variables for the 8 sites and the parameter vector $\boldsymbol{\beta}$ contains the respective coefficients.

model	formula
linear model (2)	$\eta_{ijl} = \mathbf{x}_{ijl}^T \boldsymbol{\beta} + b_i$
best model (5)	$\eta_{ijl} = \alpha + \beta_1 \text{study}_i + \text{site}_i^T \boldsymbol{\beta}_2 + f(\text{cumdos.site}_{ijl}) + b_i$
reduced model (3)	$\eta_{ijl} = \mathbf{x}_{ijl}^T \boldsymbol{\beta} + f(\text{cumdos.site}_{ijl}) + b_i$
full model (4)	$\eta_{ijl} = \mathbf{x}_{ijl}^T \boldsymbol{\beta} + f_l(\text{cumdos.site}_{ijl}) + b_i$

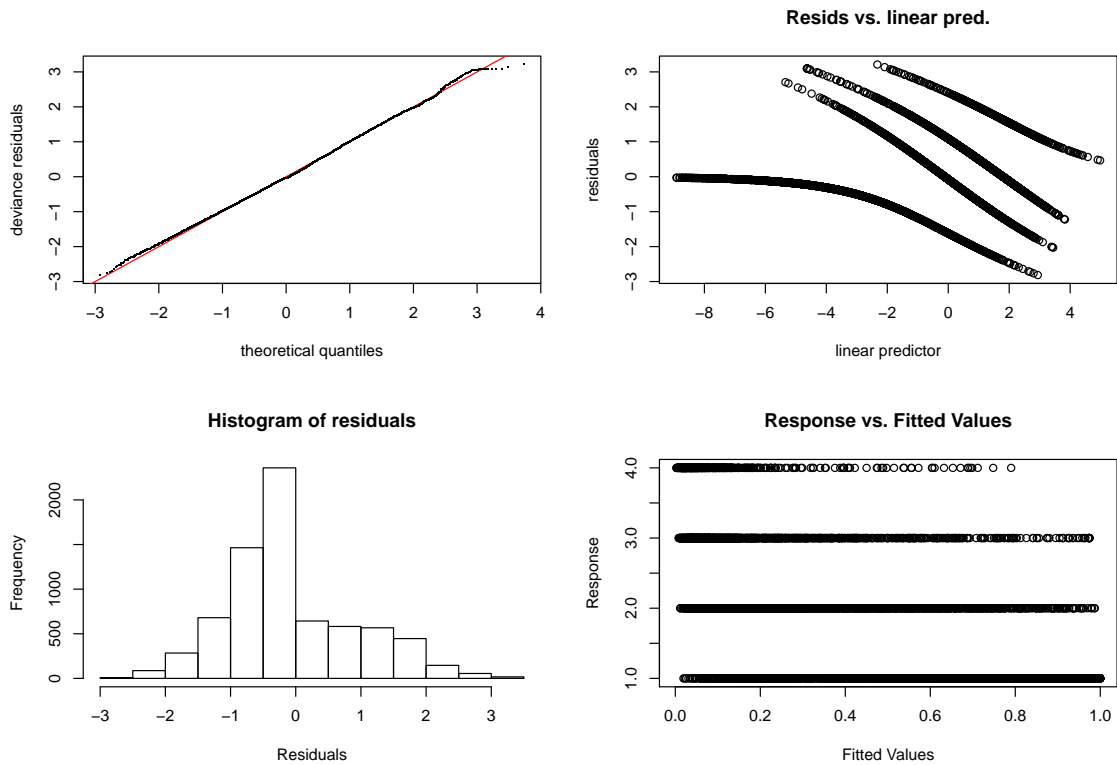


Figure 5: Residual plots of the best model (5).

Table 3: Model selection results. Shown are the effective degrees of freedom (edf), Akaike information criterion (AIC), Bayes information criterion (BIC), root mean squared error for predicting mucositis score 0 (RPE0), ... , 3 (RPE3) on the probability scale.

	N	edf	AIC	BIC	RPE0	RPE1	RPE2	RPE3
full model (4)	7339	121.92	12185.50	13085.60	0.367	0.397	0.373	0.265
reduced model (3)	7339	84.71	12380.21	12989.66	0.371	0.398	0.374	0.265
linear model (2)	7339	78.38	13076.64	13639.67	0.380	0.401	0.375	0.272
linear model (2)								
- cumdos.site	7339	77.76	15280.12	15835.42	0.455	0.407	0.390	0.278
reduced model (3) - site	7339	76.71	13284.12	13836.23	0.392	0.404	0.377	0.269
reduced model (3) - study	7339	85.05	12380.37	12992.07	0.373	0.398	0.374	0.267
reduced model (3) - age	7339	84.60	12380.17	12988.85	0.371	0.398	0.374	0.265
reduced model (3) - sex	7339	84.60	12380.17	12988.86	0.370	0.398	0.374	0.265
reduced model (3)								
- sex - age	7339	84.48	12380.13	12988.03	0.370	0.398	0.374	0.265
reduced model (3)								
- sex - age - study	7339	84.85	12380.29	12990.63	0.372	0.397	0.373	0.266

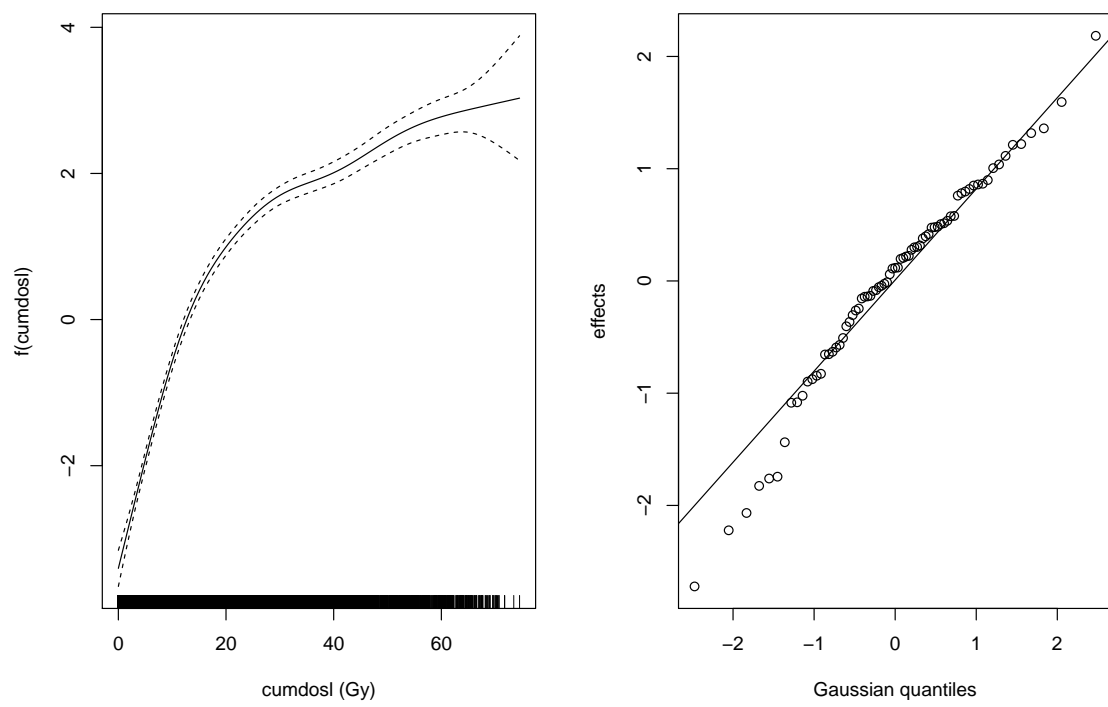


Figure 6: Best model (5): Smooth (left) and qq-plot of the random intercepts for individuals (right).

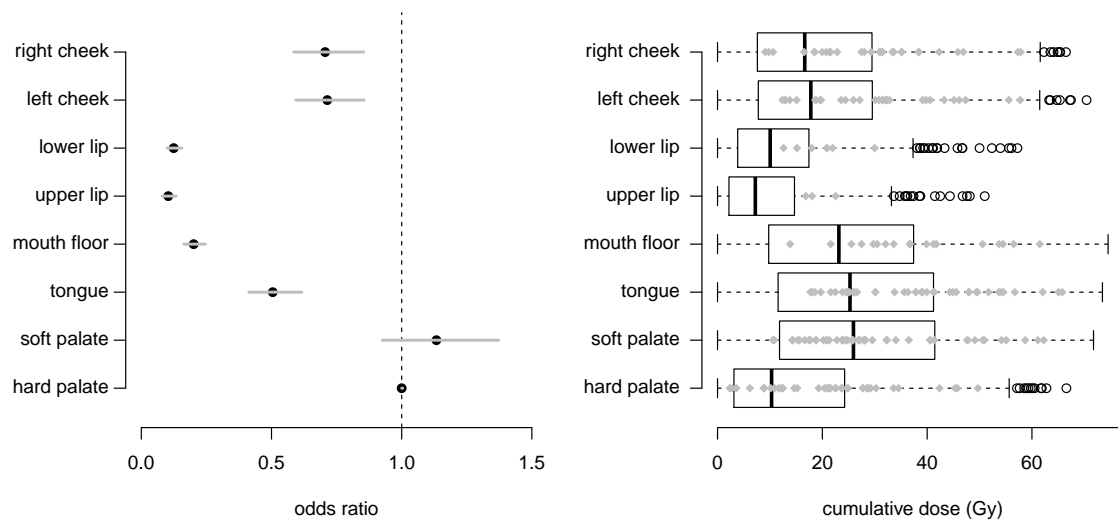


Figure 7: Best model (5). Odds ratios of all the sites with 95% confidence/credible intervals (left) and corresponding box-plots of cumulative doses applied to the eight sites (right). They are marked with gray dots by the cumulative doses when mucositis 3 or 4 first occurred.

6 Discussion

In this observational study the radiotherapists made an effort to prevent much damage of mucositis and randomized assignment of radiation dose for each site was not feasible. The assigned dose has been carefully planned to avoid severe mucositis especially for what the practitioners thought are the most vulnerable sites, upper and lower lip. Figure 2 confirms that the distribution of cumulative dose is not balanced among sites. In particular lower and upper lip and hard palate obtain the lowest median cumulative doses of 7.2 to 10.3 Gy. The highest median doses of 23.1 to 25.9 Gy arrive at soft palate, tongue and mouth floor. The left and right cheek have median doses of 16.6 to 17.8 Gy. When we investigate the frequencies of the higher non-zero mucositis scores by site in Figure 2, the same order is reflected. Upper and lower lip and hard palate have the lowest frequencies of mucositis scores and soft palate receives the highest frequencies of mucositis scores. If the dose was assigned to each site for each individual randomly, it would have been possible to estimate the site effect without adjusting for (cumulative) dose. However, the assigned dose has been carefully planned to avoid severe mucositis especially for the most vulnerable sites, that is the sensitivity of sites and site specific cumulative dose are confounded. For estimating the sensitivity of sites we need to adjust for cumulative dose in a statistical model. The selected best model (5) contains site specific cumulative dose as a non-linear effect, and the results show that after adjusting for site specific cumulative dose we find that sensitivity to radiation varies by site. The upper and lower lip are, after adjusting for cumulative dose, the least vulnerable to radiation, followed by mouth floor (Figure 7). More vulnerable than the latter are the sites tongue, right and left cheek and most vulnerable to radiation are soft and hard palate. We note that the ranking of sensitivity of sites is *not* in the same order than if we just consider marginal frequencies of mucositis scores (Figure 2), due to the confounding of cumulative dose with sites. On the boxplots of cumulative dose by site in Figure 7 we have marked the cumulative dose when mucositis score 3 or 4 first occurred. This shows that for upper and lower lip mucositis occurs rarely and at doses higher than those where it occurs at other sites. We can also see that there is a good overlap in cumulative dose when mucositis first occurred between sites.

Figure 8 compares the odds ratios of all sites from the best model (5) with the odds ratios

from logistic regression models with the same terms as the best model (5), but having collapsed the mucositis scores into a binary response using different values of r . For $r = 0$ the results are similar, but the logistic regression model yields wider confidence intervals, for $r = 1$ and $r = 2$ the confidence intervals from the logistic regression deviate substantially, and the answers are not quite as clear cut as the confidence intervals overlap. This demonstrates that we loose efficiency by aggregation.

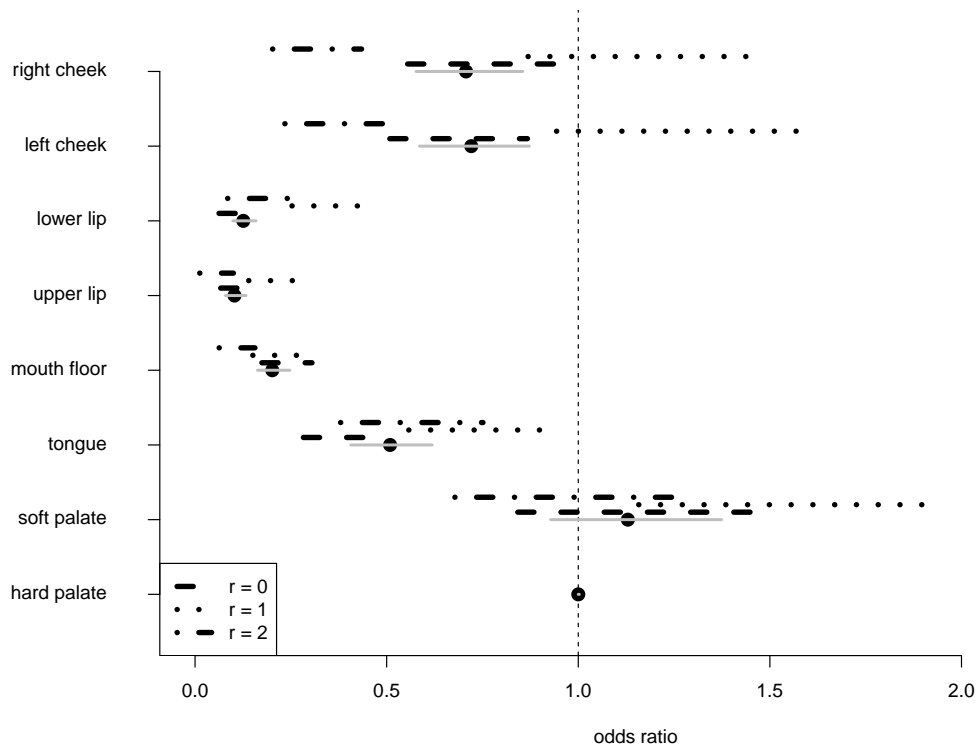


Figure 8: Odds ratios of all the sites with 95% confidence/credible intervals from the proportional odds model (best model (5)) in gray. Overlaid are the corresponding confidence intervals of odds ratios from logistic regression models with the same terms as the best model (5) having collapsed the mucositis scores into a binary response using different values of $r = 0, 1, 2$ for the two events $Y \leq r$ and $Y > r$ as shown in Table 1. Note that we do not use $r = 3$, because in the analysis we combined mucositis scores 3 and 4.

Our results depend on the accurate measurement of percentage median dose arriving at individual sites throughout the radiation schedule. It is also worth noting that we have not adjusted for

the absolute volume of the sites in our model, but we assume that the volume does not differ substantially between sites. It is future work to investigate whether site volume has an effect on mucositis. In our analysis we have used cumulative dose rather than biological effective dose, as the cumulative dose achieved the best model fit.

As in any statistical investigation, inference relies on the validity of the model. We have taken great care to avoid model misspecification by using a flexible multivariate random effects proportional odds model to estimate the effects. The model is complex, but this level of complexity is required in order to reflect the data generating mechanism. We have carried out a thorough model selection based on the AIC, BIC and the root mean square prediction error estimated by 15-fold cross-validation. The model deals with the repeated multivariate ordered categorical outcome measures of individuals via random intercepts for individuals. The model also has shown that the effect of cumulative dose is not linear, and that there is a possible interaction between site and site specific dose, as we could not entirely reject the full model (4).

Acknowledgements

This work was supported by funding from the European Communitys Seventh Framework Programme FP7/2011: Marie Curie Initial Training Network MEDIASRES (Novel Statistical Methodology for Diagnostic/Prognostic and Therapeutic Studies and Systematic Reviews; www.mediasres-itn.eu) with the Grant Agreement Number 290025.

References

- [1] Agresti, A. (2014). *Categorical data analysis*. John Wiley & Sons.
- [2] Augustin, N. H., Lang, S., Musio, M., and Von Wilpert, K. (2007). A spatial model for the needle losses of pine-trees in the forests of Baden-Württemberg: an application of bayesian structured additive regression. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 56(1):29–50.

- [3] Belitz, C., Brezger, A., Kneib, T., Lang, S., and N.Umlauf (2015). *BayesX: Software for Bayesian Inference in Structured Additive Regression Models. Version 1.0.*
- [4] Bender, R. and Grouven, U. (1997). Ordinal logistic regression in medical research. *Journal of the Royal College of Physicians of London*, 31(5):546–551.
- [5] Fahrmeir, L., Kneib, T., and Lang, S. (2004). Penalized structured additive regression for space-time data: a bayesian perspective. *Statistica Sinica*, 14:731–761.
- [6] Fahrmeir, L., Kneib, T., Lang, S., and Marx, B. (2013). *Regression: Models, methods and applications*. Springer.
- [7] Faraway, J. J. (2006). *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models*. Boca Raton, CRC press.
- [8] Fowler, J. F. (2010). 21 years of biologically effective dose. *The British Journal of Radiology*, 83:554568.
- [9] Greven, S. and Kneib, T. (2010). On the behaviour of marginal and conditional aic in linear mixed models. *Biometrika*, 97(4):773–789.
- [10] Kneib, T. and Fahrmeir, L. (2006). Structured additive regression for categorical space–time data: A mixed model approach. *Biometrics*, 62(1):109–118.
- [11] McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 109–142.
- [12] R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [13] Sakellari, I., Angelopoulou, M., Tsopra, O., Dervenoulas, I., Tsirigotis, P., Spyridonidis, A., Liga, M., Tsionos, K., Anargyrou, K., Pouli, A., and Anagnostopoulos, A. (2015). A prospective study of incidence, clinical and quality of life consequences of oral mucositis post palifermin prophylaxis in patients undergoing high-dose chemotherapy and autologous hematopoietic cell transplantation. *Annals of Hematology*, pages 1–8.

- [14] Umlauf, N., Adler, D., Kneib, T., Lang, S., and Zeileis, A. (2015). Structured additive regression models: An R interface to BayesX. *Journal of Statistical Software*, 63(21):1–46.
- [15] Verbeke, G., Fieuws, S., Molenberghs, G., and Davidian, M. (2014). The analysis of multivariate longitudinal data: a review. *Statistical methods in medical research*, 23(1):42–59.
- [16] Wood, S. (2006). *Generalized Additive Models. An Introduction with R*. Chapman & Hall/CRC, Boca Raton.
- [17] Wood, S., Pya, N., and Säfken, B. (2015). On smooth modelling with regular likelihoods. *under revision*.
- [18] Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1):3–36.